

Estils

Tecnologia excloent: quan les màquines ens discriminen

Masclistes, sexistes, homòfobes...
Les màquines inanimades que hi ha darrere de la tecnologia també tenen prejudicis

NEREIDA CARRILLO
BARCELONA

Un exercici ràpid mostra que la tecnologia no només no és infal·lible, sinó que pot arribar a ser masclista. Si busquem a Google Imatges la paraula *CEO*, no trobarem cap dona fins a la dissetena fotografia i només cinc en els 50 primers resultats. No és l'única prova de la misogínia del popular cercador. Un estudi de la Universitat Carnegie Mellon constata que Google mostra molt més als homes que no pas a les dones ofertes de feina executiva ben remunerada. La discriminació tecnològica abraça també la raça, la religió i l'orientació sexual. Per què les màquines, inanimades i insensibles, també rebutgen i marginen?

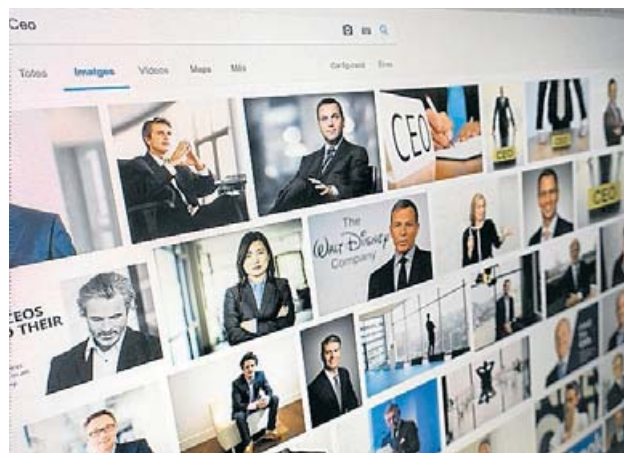
Enric Puig Punyet, filòsof i impulsor de l'associació Institut Internet, lamenta que la discriminació del vell món s'hagi escolat a les noves eines, però també que hi hagi noves maneres d'excloure. Algunes d'elles són les que emanen dels algoritmes: "Abans la discriminació es podia detectar. El problema dels algoritmes és la seva aura de misteri. Se'ns amaga la càrrega ideològica que hi ha en el fons". Que els algoritmes ens marginin no és una qüestió de poca importància. La nostra vida diària depèn de decisions que prenen les màquines. Elles determinen quina ruta agafarem, quines persones de l'apli de lligar coneixe-

rem i, fins i tot, si ens concediran un crèdit o ens acceptarà una asseguradora. "És important garantir que les decisions es prendran sense biaixos que van contra els drets humans", assenyala el director de la Càtedra Unesco de Privadesa de Dades a la URV, Josep Domingo Ferrer.

"Sabem que les decisions humanes poden ser subjectives, però esperem que les decisions preses per les màquines siguin millors, perquè es basen en dades objectives", explica Sara Hajian, analista de dades al centre tecnològic Eurecat. Diversos estudis, però, han demostrat el contrari, i alguns investigadors han començat a posar-hi remei. A Eurecat, juntament amb investigadors de la UPF i la Universitat Tècnica de Berlín, han creat un algoritme per evitar discriminacions en la cerca de persones *online*. L'eina permet reorganitzar els resultats de manera que s'eviti la marginació sense afectar la

A la imatge inferior es pot veure com les imatges d'homes són majoritàries si busquem la paraula *CEO* a internet.

RUTH MARIGOT



validesa del rànquing. Al grup de recerca encapçalat per Domingo també han dissenyat algoritmes de correcció que eliminen els biaixos.

Una marginació històrica

Els algoritmes funcionen analitzant molta informació, les anomenades dades d'entrenament, a partir de les quals extreuen conclusions i les regles que s'aplicaran. El problema és que l'algoritme pot fer encreuaments indesitjables, com ara que els negres no tornen els crèdits. "Quan entrenes un algoritme per prendre decisions, l'algoritme no sap quines variables són políticament correctes. El que fa és mirar de trobar dreceres per encertar-la al màxim", explica aquest professor de la URV. "La discriminació no és intencionada, però és la conseqüència de tot el procés", aclareix Hajian. La investigadora afegeix que sovint la marginació s'origina perquè les dades de què beuen els algoritmes reflecteixen una discriminació històrica. Si les discriminacions del passat guien les decisions del present, la marginació es perpetua. "Cal depurar les dades d'entrenament de manera que ens assegurem que no s'aprendran regles basades en atributs sensibles", proposa Domingo. Afegeix que un algoritme amb biaixos, encara que sigui molt eficaç, "no és legítim".

De vegades la discriminació de les màquines és indirecta, es produeix per variables aparentment neutres com el codi postal o el nom. És el que va passar a Airbnb. La Universitat de Harvard va demostrar que

els usuaris amb cognoms afroamericans tenien menys possibilitats de ser acceptats com a hostes que els de cognoms blancs. Comportaments com aquest van propiciar una ferotge campanya a les xarxes amb el *hashtag* #AirbnbWhileBlack. Alguns tuitaires negres denunciaven, per exemple, que quan llogaven apartaments els veïns trucaven a la policia perquè creien que hi entraven lladres. També s'han detectat comportaments racistes en algoritmes de predicció de delictes o a Flickr, que als inicis de la seva eina de reconeixement d'imatges etiquetava els negres com a animals.

Igualtat o secret industrial

Tant les marginacions directes com indirectes són invisibles en la societat. Les noves formes d'exclusió del segle XXI són més difícils de copsar i aturar. Cap banc dirà al client que no se li concedeix el crèdit per la seva raça o religió. Ni ensenyarà com funciona el seu algoritme. Ho consideren secret industrial. En alguns casos les empreses rectifiquen i prometen accions per no perdre la confiança dels usuaris, clau per al seu negoci. És el que va fer Airbnb, que va anunciar que se sotmetria a un examen del govern sobre discriminació racial. També YouTube es va disculpar i va rectificar el seu algoritme quan alguns membres de la comunitat LGTBI van detectar que els seus vídeos no apareixien en el mode restringit. Per atendre la diversitat de la societat, els emojis han anat incloent la representació de diverses

GETTY

Cal parar atenció a les dades que mostrem

.....

Els experts lamenten la vulnerabilitat dels ciutadans. S’hi pot fer poc, però per començar cal saber que les màquines no són neutres. “No hem de creure en la infal·libilitat dels algoritmes i hem de reclamar que no tot es transformi en una qüestió tecnològica”, explica el filòsof Enric Puig Punyet. Per a l’analista de dades d’Eurecat Sara Hajian, cal ser conscient dels usos que es poden donar a les dades que revelem i aconsella que no en mostrem massa. “No m’importa, per exemple, que la gent sàpiga de quin país soc; però no vull que ningú prengui la decisió de no contractar-me basada en el meu país d’origen”, afirma Hajian. Les empreses no només disposen de la informació personal que els facilitem o la que és accessible a la xarxa, sinó també de tot tipus de dades compilades pels anomenats *data brokers*, dedicats a la venda d’informació. Algunes empreses adquireixen el seu material per entrenar millor els algoritmes i que l’encertin més.

racas i religions. Però malgrat aquest tipus d’iniciatives, els experts coincideixen que els biaixos continuen sent-hi.

Les lleis espanyoles ja prohibeixen la discriminació per raons com ara el gènere, la religió o les idees polítiques. En la pràctica, en el cas dels algoritmes, com explica l’advocat i professor de dret i ciència política de la UOC Pere Vidal, resulta difícil denunciar-ho: “S’ha d’aportar una pericial informàtica que acrediti que l’algoritme utilitza aquest tipus de paràmetres. És complicat aportar aquests indicis”. Els experts reclamen a l’administració que posi el focus en el *big data*. “Per reconciliar el secret industrial amb la transparència caldrà que hi hagi una autoritat interposada”, assegura Domingo. Defensa que l’administració tingui accés als budells dels algoritmes per vetllar per la igualtat. Per a Hajian, calen solucions tant tècniques com polítiques i jurídiques: “Necessitem construir tecnologia que estigui dissenyada específicament per evitar la discriminació”.

El nou reglament europeu de protecció de dades, que entrarà en vigor el maig del 2018, prohibeix específicament discriminacions en els processos de presa de decisions automatitzats. Els experts estan a l’expectativa de com s’aplicarà, de com vetllarà l’administració per la igualtat en el *big data* i de si les empreses esmenaran els biaixos. Mentrestant, l’usuari pot fer poc més que ser conscient de l’animadversió que de vegades li professen les màquines. —